STUDIES OF MENTAL FATIGUE

W. H. HECK

Studies of Mental Fatigue

In Relation to the Daily School Program

W. H. HECK, M. A.

PROFESSOR OF EDUCATION, UNIVERSITY OF VIRGINIA

AUTHOR OF

"MENTAL DISCIPLINE AND EDUCATIONAL VALUES"

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN COLUMBIA UNIVERSITY

LYNCHBURG, VIRGINIA
J. P. BELL COMPANY, INCORPORATED, PRINTERS
1914

exch

TABLE OF CONTENTS

STUDIES OF MENTAL FATIGUE

INTRODUCTION

American books on school management and hygiene[1] show a practical though not complete agreement as to the curves of fatigue in relation to the daily school program. Conclusions from German experiments on fatigue are the main bases for this American opinion, which is somewhat as follows: The child's efficiency, on the average, is greatest from about 9:30 (after a period for settling down to work) to 11 A. M., and then declines until the noon intermission; the child comes back in the afternoon partially refreshed by the long recess and the mid-day meal, depending of course upon the length of the recess (usually one hour) and the quantity and quality of the food eaten; the second high plateau of efficiency, which is lower than the corresponding plateau of the morning session, extends from about 1:30 to 2:30 P. M., and is followed by a decline to the lowest point of the day at the time of dismissal. Some writers believe that the second plateau occurs later in the afternoon than 1:30 P. M. In a one-session school day the decline from the high morning plateau is interrupted to only a small degree by one or two short recesses, and extends to a point below that at the close of the morning session of a two-session day.

Such text-book opinions are fast becoming guides for making daily school programs throughout the United States. University departments of education, normal schools, summer schools, and teachers' reading circles are also urging reform in programs according to the following suggestions: The most difficult subjects in the curriculum, if they can be determined for a particular class, should be placed, both for class recitation and for individual study, at the periods of greatest efficiency, with special emphasis upon the high morning plateau for arithmetic

[1]For example: Arnold, *School and Class Management*, Vol. II, p. 33; Bagley, *Educative Process*, pp. 340-2, and *Classroom Management*, p. 57; Bolton, *Principles of Education*, pp. 274-5; Chancellor, *Class Teaching and Management*, pp. 122, 3; Colgrove, *The Teacher and the School*, p. 180; Cornell, *Health and Medical Inspection of School Children*, p. 193; Hollister, *High School Administration*, p. 254; O'Shea, *Dynamic Factors in Education*, pp. 286-290; Perry, *Management of a City School*, pp. 99, 100; Roark, *Economy in Education*, pp. 65, 94; Seeley, *New School Management*, p. 49; Shaw, *School Hygiene*, pp. 230, 1.

or language drill; the less difficult subjects should be distributed appropriately along the curves of fatigue; and recesses, relaxation exercises, and alternations of recitation with study periods should be used to defer and lessen the decrease in efficiency. Thus the waste of the child's energy would be checked and more school work be accomplished in a given time.

The purpose of this monograph is not to review the vast and bewildering literature on fatigue, or even that which has dominated American thought on the daily school program, but to report experiments testing the validity of the prevailing American opinion and the suggestions based thereon.[1]

---

[1]I am greatly indebted to Dr. Edward L. Thorndike, Teachers' College, Columbia University, for many suggestions regarding this work. The summaries of experiments on Fatigue in his *Educational Psychology,* vol. iii, and in Offner's *Mental Fatigue* (English translation by Whipple) render unnecessary any bibliographical discussion in this monograph. My second and third experiments were reported in the *Psychological Clinic* and the *Journal of Educational Psychology.*

FIRST STUDY

CONDITIONS.

The tests were given during the session 1911-12 to forty classes, 1,153 children, in four New York City schools, representing different nationalities and different grades of social and hygienic opportunities. On four days (each being four or five days apart) in December, 1911, four classes of fifth-grade girls, mainly Russian Jews, were tested in Public School 177; on four Fridays in January, 1912, eight classes of 5A boys, and on four Mondays eight classes of 5B boys, mainly Italians and Polish Jews, were tested in Public School 83; on four Wednesdays in January four fourth-grade and four fifth-grade boys and girls, mainly Irish, German, and American, were tested in Public School 27; on two Wednesdays in March and two in April six fifth-grade and six sixth-grade boys and girls, mainly American (with several negroes), were tested in Public School 3, Brooklyn. Fifteen of the forty classes tested were 5A classes and fifteen were 5B classes. Together they furnished the standard for the work, but two 4A, two 4B, three 6A, and three 6B classes were chosen to represent other stages of advancement. The normal-aged and over-aged (E) sections of the half-grades in P. S. 177 and the high (a), middle (b), and low (c) sections of the half-grades in P. S. 3 were included, as were also the undifferentiated sections in P. S. 83 and 27. Of course the different ability of classes of different stages of advancement did not affect my results, because in any comparison of the work at the four periods of the school day the work of the same class would be represented at each period, and thus would be compared with itself and not with a less or more advanced class. The tests in P. S. 177, 83, and 27 came before the mid-session promotion; those in P. S. 3 came after, and therefore were given to somewhat younger and less advanced children according to grade. Though more boys than girls took all four of the tests, the results fairly represent both sexes. The December, March, and April tests were given in mild weather, the January tests in cold weather; and the ventilation of the classrooms was correspondingly varied, especially by the use of window ventilation. The ventilating and lighting systems were excellent in P. S. 27, good in P. S. 177, and fair in P. S. 83 and 3. Twenty classes were tested on Wednesdays, eight on Fridays, eight on Mondays, and four on different days in the week, thus representing supposedly varied degrees of freshness in children in relation to the week's schooling. The children tested were

under medical supervision and their most obvious physical defects, especially of a contagious kind, had been looked after; but very few children had been given a general routine examination by the school doctors. The daily school programs were not uniform, sometimes even in the same school, but showed a general tendency to give arithmetic in the early morning. In P. S. 83 a group system was used throughout, about half of a class being grouped around the teacher for study while the other half carried out previous directions as to work at blackboards or desks.

The differences as to social and hygienic condition of children and schools, and as to months and days, did not affect my comparisons of the work of the classes at four periods of the school day, because the classes were tested and compared in groups of four, every class in a group having almost the same conditions and being tested on the same days. These variations in the groups as a whole were purposely sought to see if any consequent changes could be noticed in the work of the groups as compared with each other, but they were so slight as to be negligible. The differences in children and schools were representative of New York City school conditions, though the four schools selected were superior in supervision.

MATERIAL.

As my aim was to find out what decrease in efficiency resulted from the progress of the day and especially of the school work, I determined to test a large number of children at four different periods of the school day and compare the quantity and the quality of the work done. Not the slightest change was to be made in the ordinary school routine except the interruptions for the short time necessary to give the tests. My desire was to select test material as nearly like the actual lessons of the children as possible and thus avoid the artificiality, at least for my purpose, of most of the physiological and psychological tests so far made. I therefore chose, as most suitable for my purpose, the four fundamental operations of addition, subtraction, multiplication, and division with whole numbers. Of course these operations could not represent all of arithmetic or all of the curriculum, but neither could any other test material; and as the basis of all number work they were as representative as any part of the elementary curriculum could be.

In order to carry out my plan of testing children at four periods of the school day, and comparing the work done, I had

to have four different tests so similar in quantity and quality that a certain proportion of correct and incorrect work in the first test would be as nearly equal as possible to the same proportion in any one of the other three tests. Only in this way could the work at the four periods be conveniently compared. Of course it was impossible to have four different tests exactly equal as to quantity and quality, but the reader can see from the third and fourth tests herein printed that this result was very nearly approached.

As the Standard Tests in Arithmetic, by Mr. S. A. Courtis of Detroit, are widely recognized as well planned and graded, I obtained Mr. Courtis' kind permission to modify and use his Test No. 7, and also his imitation of it, which differed almost entirely by a shifting of the figures in each example. With my first and second tests thus furnished, I made out the third and fourth in imitation of these by a similar shifting of the figures. In the few examples where a figure was substituted by Mr. Courtis in the second test for one in the first, that figure reappeared in the fourth test but not in the third. Mr. Courtis' tests were not entirely suitable for my purpose in three particulars: (1) The required transfer by the child of the examples and of the answers could not be reduced to a proportionate measurement along with the work in the operations themselves. Consequently each example in my tests was printed in workable form with space underneath for the figures to be made by the child. (2) The arithmetical symbols were eliminated, and the operations necessary were designated by the words, "Add," etc., above the examples. (3) Example 1a, 2a, 9, and 17 were omitted, thus making each test consist of three examples each in addition, subtraction, multiplication, and division. The examples were printed in three parallel rows, each row with one example in each operation in the order just given. (These modifications were made by me without consulting Mr. Courtis, and he is in no way responsible for them.) The first row contained examples with no, the second with simple, and the third with advanced, borrowing and carrying. However advantageous it might have seemed for all the examples to have been of approximately equal difficulty, the wisdom of Mr. Courtis' three grades of difficulty was shown in the zest gained by the children in my tests from the rapid start with the easy examples of the first row. (The third and fourth tests are herewith given, the original size—8½ x 11 inches—and type being reduced.)

THIRD TEST.

(1) ADD 1 2 3 9012 204 2030	(2) SUBTRACT 946732 812021	(3) MULTIPLY 1202 132	(4) DIVIDE 321) 645852
(5) ADD 7403 465 5200 8001 2030 756	(6) SUBTRACT 81014232 47637698	(7) MULTIPLY 40658 564	(8) DIVIDE 436) 3297904
(9) ADD 56996 76356 68665 59877 77563 58948 35869 75883	(10) SUBTRACT 15264354 8678575	(11) MULTIPLY 68597 876	(12) DIVIDE 976) 7707472

FOURTH TEST.

(1) ADD 132 7021 301 4040	(2) SUBTRACT 974623 801212	(3) MULTIPLY 2201 312	(4) DIVIDE 231) 485562
(5) ADD 6031 754 4600 5008 6020 543	(6) SUBTRACT 74321012 47983676	(7) MULTIPLY 54068 465	(8) DIVIDE 346) 2336884
(9) ADD 89338 69866 35768 97687 39886 68985 74587 55696	(10) SUBTRACT 15465153 6587875	(11) MULTIPLY 85976 687	(12) DIVIDE 796) 5561652

METHOD.

I explained to the teachers beforehand the purpose and method of my experiments and asked them to remain in the room while the tests were being given. They were earnestly requested not to take any part in explaining or directing the work, not to discuss the tests with the children at other times, and not to vary the prescribed arithmetic lessons in the slightest so as to stress the four fundamental operations. (The children were doing more advanced work.) As far as I know, these requests were faithfully carried out, and the teachers were always courteous in making way for my work. I eliminated the teachers from any part in conducting the tests, because uniformity in method was possible only with one conductor throughout.

The presence of the teacher in a classroom gave the customary organization as a background for my tests. The method of presenting the tests to the children was worked out with great care. When I entered a classroom for the first time, the teacher directed the children to put aside whatever work they were doing individually or collectively. I then explained the nature of my four tests, the order in which the examples were to be done (with a caution not to skip the examples in division), and the directions for starting and stopping work immediately on signal. These explanations were almost the same to all classes and seemed so clear that most of the children began the first test with nearly as much understanding as they did the other three. An inevitable initial interest partly balanced whatever initial disadvantage the children may have had in the first test. The introductions to the later tests were very brief. The children thought they were being examined for correctness in speed and never realized the main purpose of the tests. Moreover, they did not know that their results would have no effect upon their school standing.

When all pencils were ready (in P. S. 83 pens were used) the teacher and I put on each desk one test sheet with the back side up. The children were not allowed to touch the tests until told to write on the back of the sheets the number of the school, the date, and the name, age, and class of pupil. (Later on I inserted the time of day when the test began.) On signal, the children turned the papers over immediately and began to work the examples in the numbered order. They worked steadily as directed, with surprisingly little hurry or interruption, until

told to stop, when they turned their papers over without putting down another figure. During the test, I quietly watched the children and their papers, so as to prevent cheating, toward which there was little inclination, and to see that the examples were worked in proper order and without hurry, especially toward the end. The children did not know how much time was being allowed, the progress of time was not mentioned, and I consulted my watch almost without being noticed.

The time limit for each test was ten minutes, not long enough for the children to be fatigued by the test itself. Great care was taken to be exact in starting and stopping. The first four (those in P. S. 177) of the forty classes tested were given twelve minutes, which time, of course, was maintained throughout the four tests with these classes, and did not affect the total results as far as my purpose was concerned—the twelve-minute tests being compared only with each other. But as many children in these four classes finished in ten or eleven minutes, the time limit was reduced to ten with the other classes. Even then several children finished a little before ten minutes and generally spent the extra time in looking over and correcting their work.

Those children that finished early put themselves at a possible disadvantage, if, in their speed, they sacrificed quality to quantity and were unable to increase the quantity further by working extra examples up to the full time allowed. However, as the total tests showed slightly more work and a slightly greater per cent of error at the periods after the first period, these disadvantages might have had slightly more influence and been added to the fatigue effect at those periods. That this difference is small is shown by the fact that the five groups which included far more than half of the completed tests had, as averages for the five groups and for the three periods, a smaller increase in amount of 0.47 per cent, and only a greater decrease in correct work of 0.04 per cent than did all the ten groups combined into one total group. If any possible disadvantage affected the results of these five groups, it was swallowed up in their superior ability. In a note on page 21 are given the comparisons of the four periods, after the results of the twelve-minute tests in P. S. 177, which many children finished, were subtracted from the total group with its similar comparisons. The comparisons, before and after this subtraction, are so nearly alike and the slight differences are so con-

flicting, as signs of fatigue, that it matters little whether the differences be attributed to early finishing or to chance variation in P. S. 177. Whatever small effect early finishing in the ten groups tested may have had upon my results, this effect would strengthen my conclusions as to the little fatigue shown in my tests, because the subtraction of it from the fatigue effect would lessen, of course, the signs of fatigue just to that degree. In fact, I might emphasize the advantage to my conclusions of the possible disadvantage of early finishing.

The children seemed to enjoy the tests. I was surprised at my success in keeping their interest nearly the same at all four periods of the school day, so that their zeal in the work would be as constant as possible and not in favor of any period. In this way there was eliminated almost entirely the element of boredom—probably the greatest influence in decreasing the quantity and quality of work in the ordinary routine of the school day. Studies of mental fatigue should be well guarded against the interference by boredom, which is often confused with fatigue, and which has vitiated the results of several experiments. This confusion is general in teachers' estimates of fatigue and renders them of little value for scientific study.

The time schedule[1] was arranged so as to test the children at approximately 9:10 A. M., 11:05 A. M., 1:10 P. M., and 2:30 P. M.—near the opening (after a brief period for settling down to work) and the close of the morning and afternoon sessions. The exact time varied from these points according to the number of classes to be tested in a day, the time of my arrival in a classroom, and the length of preparation necessary. The classes in the different schools were combined into groups of four, either one, two, or three groups being tested in a day. With only one group in P. S. 177[2] the above schedule could be closely adhered to; but with two groups a day in P. S. 27 and 83, and three groups in P. S. 3, two classes and three classes, respectively—one from each group—were tested in immediate succession at each of the four periods of the school day. The classes of the second group followed the corresponding classes of the first group just as soon as I could get from the

[1]In most classes there were two-minute formal gymnastic drills at eleven and two o'clock. I was unable to detect any effect of these drills toward increasing or decreasing fatigue; but the eleven o'clock drill may have benefited the work at the second period to a slight degree.

[2]In this school an extra class was tested at an odd period but had no place in my experiment.

one classroom to the other and start the test with the later class. In P. S. 3 the classes of the third group followed those of the second group in a similar time relationship. With two or three groups to be tested, the fourth period was begun before 2.30 P. M. so that the work could be completed before the 3:00 P. M. dismissal. The following time schedule is typical:

Time Schedule—Public School 83—5B Classes.

Group	Class	First Test	Second Test	Third Test	Fourth Test
		Jan. 8	Jan. 15	Jan. 22	Jan. 29
1	5B1	9:12	11:06	1:04	2:24
2	5B2	9:34½	11:22½	1:20	2:39
1	5B3	11:04	1:04½	2:23	9:11½
2	5B4	11:23½	1:20	2:38½	9:26½
1	5B5	1:08½	2:25	9:05½	11:04½
2	5B6	1:29	2:41	9:22	11:22
1	5B7	2:23½	9:08½	11:05	1:08½
2	5B8	2:44½	9:26	11:21	1:25

It is evident that the time relationship of the classes of the first group to each other was almost the same as that of the classes of the second, or of the third, group to each other. The results for my purpose were not affected by thus testing the classes of one group after the corresponding classes of a preceding group, because each group constituted a unit by itself for my comparison of the four periods of the school day. In fact, the advantage of a wider representation of the morning and afternoon sessions was thereby gained. For instance, the tests with the first groups about 9:10 A. M., which might possibly have come in a few classes before the children had settled down, were balanced in the total results by the tests with the second groups about 9:27 A. M. However, a comparison of the fatigue shown by these groups gives no preference to either time for the first period.

The most important features of the time schedule were the balancing and consequent neutralization of the practice effect in the second, third, and fourth tests. Thereby the fatigue effect was left free for my comparison of the quantity and quality of the work at the four periods. This was done by having each of the four tests taken at each of the four periods by one of the four classes in a group. The first test was taken at the first period by the first class, at the second period by the second class, at the third period by the third class, and at the

fourth period by the fourth class; the second test was taken at the first period by the fourth class, at the second period by the first class, at the third period by the second class, and at the fourth period by the third class; the third test was taken at the first period by the third class, at the second period by the fourth class, at the third period by the first class, and at the fourth period by the second class; the fourth test was taken at the first period by the second class, at the second period by the third class, at the third period by the fourth class, and at the fourth period by the first class. With two groups to be tested in one day, the four classes in the second group followed the above schedule about seventeen minutes after the corresponding classes of the first group. In P. S. 3 the third group followed the second group in similar order. By such a schedule the work of the group at each period included all four tests (each test was generally taken a week apart), and consequently the practice effect of one class at the second period, of another class at the third period, and of another class at the fourth period.

This arrangement was based on the belief that, if each test were taken by each class at the same period, the practice effect of any one class in the second, third, or fourth test would be almost equal to the practice effect of any other class in its group in the corresponding test; and that, as each period included the practice effects in the second, third, and fourth tests, the sum of the three practice effects of a group at any one period would be balanced and neutralized by the sum of the three almost equal practice effects of the same group at any other period. This belief in the approximate equality of the practice effect of the four classes in a group can not be proved or disproved by my results, because the same test was taken at different periods by the four classes, and the practice effect of any class was thereby complicated with a different fatigue effect and rendered incomparable with the practice effect of any other class in its group in the same test. If there were differences in the practice effect of the four classes in a group, especially of those classes at different stages of advancement, they were probably very slight. Even these differences were partially neutralized in my total results, since the classes in different groups were not tested in the same order of advancement: in P. S. 177 in ascending order, in P. S. 83 on the same level of advancement, in P. S. 27 in descending order, and in P. S. 3 in ascending order. (As the four half-grades in P. S. 3 were divided into a, b, and c sections according to

ability, these sections were combined into a, b, and c groups; and the results show some positive though not consistent correlation between ability and resistance to fatigue.) This variation gave a different relationship between practice and fatigue effects in the different groups according to the periods at which the corresponding tests were taken by the more advanced and by the less advanced classes.

The figures on page 26 show that the practice effect in the forty classes tested was small, especially because (1) an initial interest partly balanced whatever initial disadvantage the children may have had in the first test, (2) the work was rapid and short, (3) a week generally intervened between the tests, and (4) no class drill on the subject-matter of the tests was allowed. However, the elimination of the practice effect from my results frees this experiment from the error common to most of the experiments on mental fatigue in relation to the daily school program.

The practice effect within a single test was not considered. With tests having simple examples of the same kind, such a result were possible; but with the variety of examples in my tests, both in kind and in difficulty, the accumulation of a practice effect during the ten minutes could hardly have taken place. Even if there were such practice effect in my results, it would have been balanced against itself in my comparison of the four tests at the four periods.

GRADING.

The papers were filed and graded, and the results tabulated, according to school, test, class, and name. Only full sets of four tests per child were used, all partial sets being thrown out. Unless called back to their former class for my work, the children who were removed to another class, after one or more tests, took the remaining tests at wrong periods, thus causing all their papers to be thrown out. In spite of many absences and removals, 1,153 children in the forty classes worked the four tests in proper order.

The grading of the 4,612 papers necessitated such tedious and conscientious attention to details, that it was not "farmed out" to others, who might have been careless, but was done entirely by my wife and me. The care taken to be exact and to go over the papers again in search for possible errors in grading extended the time for this work to four months. In

the detailed method used, every figure the child put down in working the examples was graded right or wrong, a complete paper in each test including 190 figures. No child was given a total credit of right or wrong figures for an example or part of an example greater than that necessary for the correct working of that example or that part, though the mistake of a child would often involve a greater number of figures. On the other hand, the completion of work on an example or part of an example gave the child the total credit of right or wrong figures for that example or that part, whether the child had put down that number of figures or not. Misplaced figures were counted wrong, as were also the omitted figures in those examples that had been worked by the child. In addition to these main rules for grading the papers, there were many subsidiary and consistent rules for special cases, which helped greatly in threading the jungle of right and wrong figures which a child would sometimes create.

Every method of grading arithmetic papers is open to objections, but, for my exact comparison of the work at four periods of the school day, the detailed method used seemed far preferable. Most of the objections which might be urged against this method would not apply to the present case, as the aim was not to give an absolute grading of ability in the four fundamental operations, but simply to be consistent in judging the quantity and quality of the work at the four periods. It might be urged that the value of the figures in some examples or parts of examples was not equal to that of those in other examples or parts of examples. But where a child worked the similar examples in all four tests the method of marking was the same for all tests, and therefore for all periods. It is only where a child worked one or more examples in one or more tests, but not the similar examples in the other tests, and where the figures in these particular examples could be shown to be undervalued in comparison with the figures in those examples that had been worked in all four tests, that the above-mentioned objection has any validity. The valuation of the figures in the ninth example in comparison with that of those in the fifth, in the few cases where the ninth example was worked by a child in some tests but not in others, is the only illustration that might possibly affect my results to more than a negligible degree. Whatever small effect a possible disparity in the value of figures in these and other examples might have had upon my results would strengthen my conclusions as to the little

fatigue shown in my tests, because the subtraction of this effect from the fatigue effect would lessen, of course, the signs of fatigue just to that degree. In fact, I might emphasize the advantage to my conclusions of this possible disparity.

Another objection may be made that an initial wrong figure may involve a series of figures which, though wrong, are really the correct consequence of the first mistake and that, therefore, only the initial figure should be counted wrong. To trace out from every initial wrong figure in every example in 4,612 papers all the consequent figures, and to determine whether the latter figures were correct consequences of the former, would be an interminable task. Even then there would never be certainty that the consequent figures, in individual papers, did not include more initial wrong figures than the first one. As the necessary thing in my experiment was to grade consistently the work at all four periods, and, as every figure was graded according to the same rules for these periods, the change suggested, even if possible, would not have affected my results.

In order to strengthen my comparison of the work at the four periods by using two different methods of grading, rather than one alone, all the papers were graded again by a gross method in which only the answers were considered. The gross method chosen was very similar to that used by Mr. Courtis. The first and second examples were together credited with one point, this point being counted wrong on account of one or more mistakes in either example; the third, fourth, fifth, sixth, and tenth examples were each credited with one point, the point for a given example being counted wrong on account of one or more mistakes in that example; the seventh, eighth, ninth, eleventh, and twelfth examples were each credited with two points, one point for a given example being counted wrong on account of one mistake, and two points being counted wrong on account of two or more mistakes in that example.

RESULTS.

After the results for each child were graded by the detailed method and tabulated, the number of right and wrong figures for each class in each test was averaged, and then the number of right and wrong figures for each group of four classes at each of the four periods of the school day. The individual variations in the children were first merged in the class aver-

ages and again in the group averages. From these averages were made tables for each of the ten groups. The following table is typical:

Table of Results—Group 2—Public School 27.

Class	Pupils	FIRST PERIOD			SECOND PERIOD			THIRD PERIOD			FOURTH PERIOD		
		Test	Right	Wrong	Test	Right	Wrong	Test	Right	Wrong	Test	Right	Wrong
5B1	26	1	129.69	14.92	2	132.35	25.31	3	142.77	23.42	4	140.31	25.27
5A1	43	4	149.07	15.98	1	126.98	12.12	2	142.16	21.44	3	145.28	20.21
4B2	32	3	131.53	17.53	4	122.72	24.84	1	109.56	22.25	2	124.50	20.47
4A2	26	2	112.08	20.65	3	126.69	24.69	4	122.96	28.35	1	101.88	23.58
4	127		130.59	17.27		127.18	21.74		129.36	23.86		127.99	22.38
Sum of right and wrong			147.86			148.92			153.22			150.37	
Relative amount done			100.00			100.72			103.63			101.70	
Per cent right			88.32			85.40			84.43			85.12	
Relative per cent right			100.00			96.69			95.64			96.38	

These tables gave answers to two questions, which formed the core of the investigation. The answer to first question, What was the difference in the quantity of work done at the four periods?, was reached by getting the sum of right and wrong figures at each period, by giving the value of 100.00 to the sum at the first period (of course any other period could have been taken as the standard of comparison), and then by calculating the per cents of this sum represented by the sum at each of the other three periods. These per cents for the ten groups were as follows, the per cent for the first period being 100.00 in every group:

Group	School	Second Period	Third Period	Fourth Period
1	177	103.16	101.36	103.19
2	27	100.72	103.63	101.70
3	27	100.14	98.28	96.62
4	83	100.49	100.37	105.81
5	83	98.85	98.60	104.05
6	83	102.42	101.65	103.52
7	83	102.15	102.17	101.87
8	3	101.41	101.20	103.74
9	3	101.05	104.40	99.99
10	3	104.34	104.26	104.54

The answer to the second question, What was the difference in the quality of work done at the four periods?, was reached by calculating the per cent right of the sum of right and wrong figures at each period, by giving the value of 100.00 to the per cent right at the first period, and then by calculating the per cent of this per cent right represented by the per cent at each of the other three periods. These per cents for the ten groups were as follows, the per cent for the first period being 100.00 in every case:

Group	School	Second Period	Third Period	Fourth Period
1	177	97.53	97.58	97.37
2	27	96.69	95.64	96.38
3	27	98.86	97.34	97.54
4	83	100.39	98.65	97.94
5	83	98.92	96.90	95.89
6	83	97.64	100.94	98.65
7	83	97.81	99.12	98.91
8	3	101.74	101.39	100.52
9	3	98.75	99.15	97.45
10	3	96.75	99.06	97.08

In *quantity* of work done (sum of right and wrong figures) as compared with the first period, seven of the ten groups tested showed an increase at three periods, one an increase at two periods and a decrease at one period, and two an increase at one period and a decrease at two periods. In all, there were twenty-five cases of increase, averaging 2.53 per cent; and five cases of decrease, averaging 1.53 per cent. The same facts by periods were as follows: at the second period an average increase of 1.76 per cent in nine groups and a decrease of 1.15 per cent in one; at the third period an average increase of 2.38 per cent in eight groups and an average decrease of 1.56 per cent in two; at the fourth period an average increase of 3.55 per cent in eight groups and an average decrease of 1.69 per cent in two.

In *quality* of work done (per cent of sum right) as compared with the first period, seven of the ten groups tested showed a decrease at three periods, two a decrease at two periods and an increase at one period, and one an increase at all periods. In all, there were twenty-five cases of decrease, averaging 2.26 per cent, and five cases of increase averaging 1.00 per cent. (The five cases of increase in quality occurred in different groups from those having the five cases of decrease in quantity; consequently there was no connection between them.) The

same facts by periods were as follows: At the second period an average decrease of 2.13 per cent in eight groups and an average increase of 1.06 per cent in two groups; at the third period an average decrease of 2.07 per cent in eight groups and an average increase of 1.16 per cent in two groups; at the fourth period an average decrease of 2.53 per cent in nine groups and an increase of 0.52 per cent in one group.[1]

The most important summary of results is shown in the total table combining the forty classes of the ten groups tested, in which individual, class, and group variations were all merged. The average of the forty class averages for each period was used as the material for calculating the comparative per cents of quantity and quality exactly in the same way that the average of four class averages for each period was used in a table for each single group.[2] The balanced relation to each other of the four classes in each group was not disturbed by this combination, which tended, however, to neutralize whatever chance variations may have occurred in the single groups. *In quantity of work done (sum of right and wrong figures) as compared with the first period of the school day, the ten groups together showed a respective increase of 1.57, 1.64, and 2.36 per cent at the second, third, and fourth periods; and in quality of work done (per cent of sum right) a respective decrease of 1.51, 1.41, and 2.22 per cent at the second, third, and fourth periods. For comparison with these results, obtained by the detailed method of grading, a similar total table was made with the class averages by the gross method of grading, and the following per cents were calculated: In quantity of work done (sum of right and wrong points) as compared with the first period, the ten groups together showed a respective increase of 1.92, 1.31, and 2.27 at the second, third, and fourth periods; and in quality of work done (per cent of sum right) a respective decrease of*

[1]In order to give also the results by schools, the first and second groups of P. S. 27, of P. S. 83-5A, and of P. S. 83-5B, and the first, second, and third groups of P. S. 3, were combined into four tables by getting the group averages for each period of the eight or twelve class averages in the groups combined, and calculating therewith the comparative per cent of quantity and quality exactly in the same way as in the single groups.

[2]The average of the forty class averages was used instead of the average of the ten group averages, as being more representative of the total conditions; but the difference between the two total averages would be very small.

5.74, 4.56, and 6.66 at the second, third, and fourth periods.[1] The per cents of increase in quantity in the two tables are surprisingly similar; the per cents of decrease in quality in the second table are a little over three times as large as those in the first table, but are still very small. This difference is probably due to the greater weight given errors by the gross method of grading, one error in the answer to any example making wrong one point out of a total of sixteen points for the entire test. This seems to me a defect natural to any gross method of grading for purposes of scientific study. (Notice the difference in the per cents of decrease in quality in the practice results on page 26.) However, it would be inevitable that any two, or any three, methods of grading would show some differences in results. The two total tables are given herewith. The one for the gross method of grading contains, of course, all the class averages necessary to make out the single table for each group; but this extra work was not needed and therefore was not undertaken.

[1]Subtracting the four classes of P. S. 177, as suggested on page 12, we have for thirty-six classes, by the detailed method of grading, an increase of 1.35, 1.68, and 2.24 per cent, respectively, in quantity, and a decrease of 1.39, 1.26, and 2.17 per cent, respectively, in quality; by the gross method of grading an increase of 1.70, 1.52, and 1.97 per cent, respectively, in quantity, and a decrease of 6.31, 4.41, and 6.59 per cent, respectively, in quality.

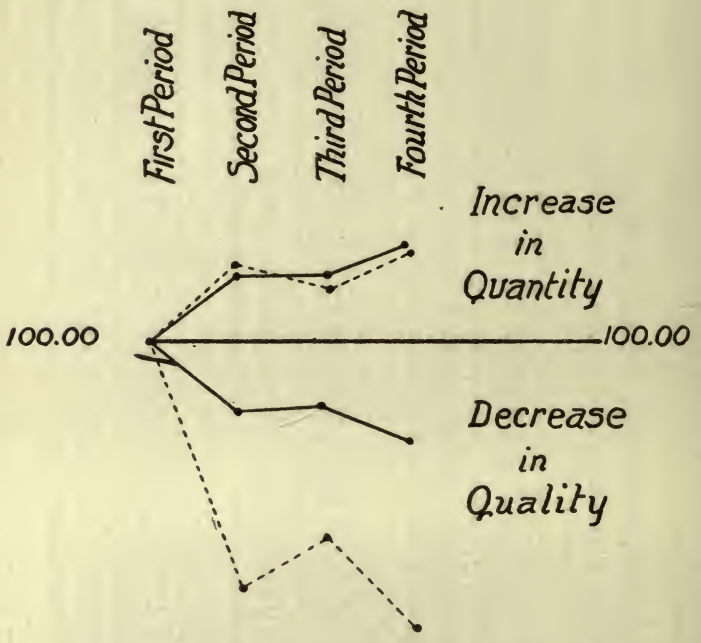
Forty Classes (Ten Groups) Combined Into One Group. Detailed Method of Grading.

Group	School	Class	Pupils	FIRST PERIOD			SECOND PERIOD			THIRD PERIOD			FOURTH PERIOD		
				Test	Right	Wrong	Test	Right	Wrong	Test	Right	Wrong	Test	Right	Wrong
1	177	5A1	29	1	150.34	10.59	2	155.28	23.17	3	162.07	16.31	4	167.79	19.14
2	27	5B1	26	1	129.27	15.35	2	131.96	25.31	3	142.77	23.42	4	140.31	25.27
3	27	5B2	25	1	135.80	15.52	2	142.76	26.56	3	144.36	27.48	4	137.36	28.56
4	83	5A1	35	1	99.09	19.00	2	108.09	16.94	3	113.26	19.60	4	116.97	28.86
5	83	5A2	37	1	92.22	19.51	2	103.57	19.19	3	109.16	21.49	4	109.57	27.24
6	83	5B1	36	1	100.47	12.69	2	110.33	16.72	3	123.42	11.28	4	119.78	17.19
7	83	5B2	32	1	114.12	9.91	2	117.41	19.91	3	126.06	14.16	4	123.78	15.06
8	3	5Aa	28	1	110.11	15.46	2	115.86	17.54	3	125.14	15.68	4	127.54	18.18
9	3	5Ab	19	1	109.89	23.74	2	115.47	23.32	3	130.37	29.16	4	112.79	30.76
10	3	5Ac	19	1	92.95	14.05	2	96.42	24.11	3	113.79	21.89	4	101.47	25.16
1	177	5B1	38	4	159.67	11.34	1	155.68	12.42	2	158.53	11.11	3	164.32	10.53
2	27	5A1	43	4	148.81	16.58	1	126.93	12.16	2	142.88	21.70	3	144.88	20.21
3	27	5A2	31	4	153.16	17.71	1	138.13	12.06	2	144.00	21.00	3	151.84	22.06
4	83	5A3	38	4	102.92	20.89	1	92.63	16.08	2	99.05	18.79	3	112.92	17.71
5	83	5A4	26	4	110.46	25.81	1	87.62	18.12	2	98.35	25.81	3	111.19	25.92
6	83	5B3	28	4	111.86	18.43	1	100.32	15.61	2	110.68	15.50	3	112.43	18.89
7	83	5B4	27	4	114.78	20.15	1	98.63	18.30	2	113.07	18.26	3	118.96	21.00
8	3	5Ba	29	4	132.14	16.93	1	126.17	15.14	2	127.59	20.66	3	133.07	17.79
9	3	5Bb	23	4	126.43	24.43	1	116.04	25.96	2	122.30	29.00	3	123.74	23.26
10	3	5Bc	15	4	106.67	28.00	1	106.13	26.60	2	113.40	28.87	3	111.20	32.60
1	177	5Ea	40	3	164.60	8.92	4	158.17	13.35	1	136.35	16.15	2	143.62	16.07
2	27	4B2	32	3	131.66	17.56	4	122.53	24.84	1	109.22	22.56	2	123.87	21.03
3	27	4B3	27	3	137.37	22.85	4	130.33	31.59	1	113.37	22.89	2	123.48	25.15
4	83	5A5	40	3	107.42	20.07	4	106.60	24.40	1	93.12	18.00	2	105.47	23.05
5	83	5A6	29	3	116.59	13.14	4	117.41	17.90	1	90.14	17.62	2	105.79	19.01
6	83	5B5	25	3	110.60	11.80	4	106.04	20.60	1	89.24	11.04	2	107.00	18.04
7	83	5B6	27	3	108.15	20.26	4	108.63	22.56	1	95.04	19.59	2	104.07	21.44
8	3	6Aa	30	3	155.97	14.47	4	156.93	12.83	1	151.90	9.97	2	152.77	18.20
9	3	6Ab	34	3	149.76	15.68	4	143.68	20.91	1	136.03	19.50	2	138.85	21.91
10	3	6Ac	24	3	126.04	21.50	4	120.29	31.00	1	112.21	24.29	2	122.12	25.54
1	177	5Eb	31	2	152.26	13.71	3	161.61	13.00	4	163.10	16.94	1	154.16	17.23
2	27	4A2	26	2	112.08	20.65	3	126.65	24.73	4	122.96	28.35	1	101.88	23.58
3	27	4A3	25	2	101.88	31.96	3	114.12	22.12	4	105.92	27.60	1	86.56	20.48
4	83	5A7	29	2	99.31	14.17	3	105.00	19.45	4	99.28	23.55	1	88.21	17.72
5	83	5A8	29	2	94.79	11.86	3	96.28	18.72	4	97.93	17.07	1	86.52	18.72
6	83	5B7	26	2	122.50	16.73	3	128.77	18.92	4	133.73	18.50	1	113.46	16.46
7	83	5B8	21	2	115.95	16.86	3	127.95	18.00	4	124.57	20.71	1	109.62	15.95
8	3	6Ba	28	2	140.07	20.07	3	156.43	12.86	4	147.64	13.86	1	147.89	12.39
9	3	6Bb	25	2	128.80	17.36	3	138.64	18.32	4	144.32	11.68	1	126.36	18.40
10	3	6Bc	21	2	130.19	21.71	3	137.38	22.67	4	131.43	18.24	1	127.86	19.71
		40	1,153		122.68	17.69		122.72	19.85		122.94	19.73		122.79	20.89
Sum of right and wrong					140.37			142.57			142.67			143.68	
Relative amount done					100.00			101.57			101.64			102.36	
Per cent right.....					87.40			86.08			86.17			85.46	
Relative per cent right					100.00			98.49			98.59			97.78	

Forty Classes (Ten Groups) Combined Into One Group. Gross Method of Grading.

Group	School	Class	Pupils	FIRST PERIOD			SECOND PERIOD			THIRD PERIOD			FOURTH PERIOD		
				Test	Right	Wrong	Test	Right	Wrong	Test	Right	Wrong	Test	Right	Wrong
1	177	5A1	29	1	11.38	2.21	2	10.69	3.90	3	11.28	3.62	4	11.66	3.79
2	27	5B1	26	1	9.23	2.96	2	8.58	4.35	3	9.27	4.54	4	9.00	4.65
3	27	5B2	25	1	9.40	3.28	2	9.20	5.00	3	8.92	5.36	4	8.48	5.20
4	83	5A1	35	1	6.14	2.89	2	6.51	3.80	3	6.91	3.71	4	7.11	4.66
5	83	5A2	37	1	6.03	3.24	2	6.38	3.43	3	6.46	4.62	4	6.49	4.84
6	83	5B1	36	1	6.56	2.22	2	7.33	2.75	3	8.44	2.47	4	7.92	3.33
7	83	5B2	32	1	7.28	2.62	2	7.28	4.16	3	8.62	3.16	4	7.84	3.87
8	3	5Aa	28	1	7.57	2.50	2	7.21	3.75	3	8.46	3.25	4	7.89	3.64
9	3	5Ab	19	1	7.11	4.32	2	7.42	4.37	3	7.89	5.32	4	5.84	5.79
10	3	5Ac	19	1	5.79	2.26	2	5.47	4.63	3	6.89	4.53	4	6.16	4.32
1	177	5B1	38	4	11.42	2.76	1	10.92	3.18	2	11.76	1.87	3	11.66	2.84
2	27	5A1	43	4	10.30	3.33	1	8.35	3.12	2	9.47	4.07	3	8.84	4.79
3	27	5A2	31	4	10.10	3.90	1	9.61	3.00	2	9.74	3.90	3	9.52	4.90
4	83	5A3	38	4	6.76	2.87	1	5.82	2.39	2	6.13	3.32	3	7.18	3.68
5	83	5A4	26	4	6.92	3.88	1	5.15	3.08	2	5.77	3.92	3	6.42	4.96
6	83	5B3	28	4	7.46	3.04	1	6.14	3.07	2	7.11	3.04	3	7.07	3.54
7	83	5B4	27	4	7.04	3.96	1	5.85	3.78	2	7.26	3.33	3	6.93	4.15
8	3	5Ba	29	4	8.59	3.48	1	7.48	3.72	2	8.28	3.83	3	8.07	4.07
9	3	5Bb	23	4	7.61	4.65	1	6.57	4.78	2	7.30	4.83	3	7.57	4.48
10	3	5Bc	15	4	6.80	4.33	1	6.73	4.27	2	6.13	5.53	3	6.53	5.27
1	177	5Ea	40	3	11.92	2.42	4	11.52	2.92	1	9.20	3.20	2	10.37	3.02
2	27	4B2	32	3	8.59	3.72	4	7.16	4.91	1	6.12	4.34	2	7.84	4.06
3	27	4B3	27	3	8.74	4.30	4	8.07	5.37	1	6.67	4.44	2	7.70	4.37
4	83	5A5	40	3	5.90	4.20	4	6.67	3.80	1	5.67	2.50	2	6.65	3.07
5	83	5A6	29	3	7.79	2.62	4	7.59	3.31	1	5.59	2.45	2	7.00	3.00
6	83	5B5	25	3	7.20	2.80	4	6.80	3.64	1	5.16	2.40	2	6.12	3.60
7	83	5B6	27	3	7.48	2.85	4	7.07	3.44	1	5.96	3.19	2	6.67	3.19
8	3	6Aa	30	3	10.77	3.43	4	10.57	3.47	1	10.97	2.40	2	10.63	3.60
9	3	6Ab	34	3	9.71	4.21	4	9.41	4.00	1	8.76	4.24	2	9.29	4.00
10	3	6Ac	24	3	8.21	4.17	4	6.96	5.29	1	7.25	3.92	2	7.00	5.00
1	177	5Eb	31	2	11.45	2.39	3	11.13	3.23	4	11.32	3.35	1	10.74	3.71
2	27	4A2	26	2	7.23	3.27	3	7.62	4.69	4	7.50	4.58	1	6.31	3.42
3	27	4A3	25	2	6.12	4.72	3	7.20	3.96	4	6.04	4.48	1	5.32	3.08
4	83	5A7	29	2	6.14	2.93	3	6.86	3.72	4	6.17	4.03	1	5.34	2.79
5	83	5A8	29	2	6.17	2.52	3	5.66	3.41	4	5.66	3.24	1	5.34	3.03
6	83	5B7	26	2	8.65	2.65	3	8.12	4.04	4	9.35	3.27	1	7.35	3.15
7	83	5B8	21	2	8.33	2.52	3	8.90	3.05	4	8.24	3.38	1	7.05	3.14
8	3	6Ba	28	2	9.39	3.79	3	10.46	3.50	4	9.64	3.36	1	10.25	2.96
9	3	6Bb	25	2	8.24	3.72	3	8.48	4.52	4	9.72	3.32	1	8.04	3.76
10	3	6Bc	21	2	8.95	3.76	3	8.62	4.48	4	8.48	3.90	1	8.43	3.95
		40	1,153		8.16	3.29		7.84	3.83		7.89	3.71		7.79	3.92
Sum of right and wrong					11.45			11.67			11.60			11.71	
Relative amount done					100.00			101.92			101.31			102.27	
Per cent right.....					71.27			67.18			68.02			66.52	
Relative per cent right					100.00			94.26			95.44			93.34	

The following diagram illustrates these comparative per cents of increase in quantity (above the heavy horizontal line) and of decrease in quality (below the heavy horizontal line). The continuous lines represent the results by the detailed method of grading, and the dotted lines the results by the gross method of grading.



In these two series of per cents are apparently opposing results—a fairly continuous increase in quantity of work done at the four successive periods, with a very small increase in the second table at the third period; and a fairly continuous decrease in quality, with a very small increase in the first table and a small increase in the second table at the third period. (These variations at the third period were due to the influence of the intermission from twelve to one o'clock.) However, this opposition may be more apparent than real. Fatigue often shows itself in a weakening of that inhibition which assures slower but more careful and consequently more correct work, the speed of the fatigued worker generally being a sign of decreased efficiency. (This result is evident in many

different experiments.) But a claim that all the above-mentioned increase in quantity is due to fatigue would be a biased interpretation, though the influence of the noon intermission in arresting the increase at the third period and the presence of the greatest increase at the fourth period strengthen this claim. It is possible that some of the increase may be due to increased efficiency as the day advances, the children becoming more habituated to school work—having a greater "swing," as Offner says. The decrease in quality of work done at the four successive periods may more justly be considered as due almost entirely to fatigue, which is arrested by the noon intermission but is greatest at the fourth period. A slight boredom and consequent carelessness as the day advances may have affected the results in spite of my efforts to keep the interest nearly the same at all four periods. Altogether, we may consider that most, though not all, of this increase in quantity and decrease in quality is a sign of fatigue.

It is impossible to combine accurately into one series for each table the per cents of quantity and of quality at each of the four periods in order to make a single comparison of the work done. Any such combination, though arbitrary, ought to be based upon a just proportion of value between the two series. For instance, one per cent of increase in quantity should not cancel one per cent of decrease in quality because the latter per cent has more value than the former, more weight in grading ability. A suggestive proportion is that based upon the per cents right of the sum of right and wrong figures (or points) at the second, third, and fourth periods, these being in the first table 86.06, 86.17, and 85.46 per cent, respectively. These per cents give a ratio of value of 1.00:.8608 between the sum and the per cent right at the second period, a ratio of 1.00:.8617 at the third period, and a ratio of 1.00:.8546 at the fourth period. For the second period we multiply the 1.57 per cent in quantity by .8608 and the 1.51 per cent in quality by 1.00, in order to reduce the per cents to equal value, and get 1.35 and 1.51 per cent, respectively. By subtracting the per cent of increase in quantity from the per cent of decrease in quality, we get a remaining decrease of .16 per cent. This gives 99.84 as the combined per cent of quantity and quality at the second period. In the same way we get 100.00 as the combined per cent at the third period and 99.80 at the fourth. The combined per cents in the second table are 95.55, 96.33, and 94.85 for the second,

third, and fourth periods, respectively. It might be objected that, as both the increase in quantity and the decrease in quality are mainly due to fatigue, they should not be balanced against each other and only the remainder be subtracted from the 100.00 per cent standard of the first period. One might even go so far as to suggest subtracting from this 100.00 per cent standard the sum of the per cents of increase and decrease. The reply is that the per cents of comparison must be based upon the actual work done, not upon an interpretation of the causes thereof, and that in grading such work an increase in quantity would naturally tend to offset a decrease in quality and vice versa, whether or not both the increase and the decrease could be attributed to one cause. However, my suggested combination may show too little fatigue and is offered simply as an illustration of a possible method of reaching one series of per cents for purposes of comparison.

The *practice effect* was calculated by rearranging the figures in the two total tables in columns according to tests rather than periods. As each test was given forty times—ten times at each of the four periods—the total result for each test combined the work at each of the four periods and thus neutralized the fatigue effect, which has often influenced the results in practice experiments. These total results were compared to find the practice effect. By the detailed method of grading, in quantity of work (sum of right and wrong figures) as compared with the first test, the ten groups together showed a respective increase of 8.91, 14.09, and 13.67 per cent in the second, third, and fourth tests; in quality of work done (per cent of sum right) a respective decrease of 1.35 and 1.31 per cent in the second and fourth tests, and an increase of 0.46 per cent in the third test. By the gross method of grading, in quantity of work done (sum of right and wrong points) as compared with the first test, the ten groups together showed a respective increase of 8.97, 15.39, and 13.97 per cent in the second, third, and fourth tests; and in quality of work done (per cent of sum right) a respective decrease of 2.00, 2.28, and 3.20 per cent in the second, third, and fourth tests.

SECOND STUDY

The tests were given in February and March, 1913, to sixteen classes, 573 pupils, in three Lynchburg, Va., schools. The important difference between this experiment and that in New York City is the amount of continuous work required by the tests. The New York classes were given four ten-minute tests at four periods of the school day, the Lynchburg classes two twenty-five minute tests at two periods.[1] The subject matter used was the same in both experiments, but the Lynchburg classes took two tests as one. The reason for this extension of time was the supposition that a ten-minute spurt might not reveal the amount of fatigue present, the pupils working the short test with fairly unform efficiency at four periods of the school day. A continuous application for twenty-five minutes could hardly fail to reveal most of the fatigue present at the time of starting the test and the further possible fatigue from the longer test itself. If the fatigue from the test itself were equally present in the morning and in the afternoon work, it would be neutralized in the results; but if it were greater in the afternoon on account of greater fatigue at the time of starting, the difference would be evident in the results. The longer test, therefore, has two possibilities of revealing fatigue and is also a better illustration of the time required, though with less continuous pressure, for study and recitation in actual school practice.

The Lynchburg tests were given on February 25 and 27, 1913, to the 5A, 5B, 6A, and 6B classes in Biggers School; on February 26 and 28 to the 5A, 5B, 6A, and 6B classes in Monroe School; on March 11 and 13 to the 7B, 7A, 6B, and 6A1 classes in Federal School; on March 12 and 14 to the 6A2, 5B1, 5B2, and 5A classes in Federal School. The weather was mild, and window ventilation was used in addition to the furnace system. The schools had a one-session day, with a fifteen-minute recess soon after the morning tests and a twenty (sometimes fifteen) minute recess before the afternoon tests. During the second recess many pupils ate lunch at home or school. This recess probably increased the efficiency in the afternoon tests; but as every school has or should have a recess near this period of the school day, the conditions of the Lynchburg tests were typical, certainly for a one-session school day. The de-

[1]As the Lynchburg classes were slower in this kind of work, twenty-five rather than twenty minutes were allowed.

partmental system of instruction allowed no uniform program of recitations and study periods in the classes tested. The children represented for the most part hygienic opportunities at home and had been examined by their teachers for eyesight and hearing. All the classes included both boys and girls, the total being 292 boys and 281 girls. The average age was 12.55 years.

The following description is condensed and adapted from the First Study, to which the reader is referred for further discussion. The departmental teacher put aside her work for my test, but remained in the room at my request. I gave detailed directions to the class and then with the help of the teacher put on each desk a test paper with the blank side up. After the name, etc., had been written, the children turned the papers over immediately on signal and worked steadily for twenty-five minutes. Great care was taken to be exact in starting and stopping. The children thought they were being examined for correctness and speed, but did not know how much time was being allowed and did not rush toward the end. The very few children who finished before the time was out looked over part of their papers. The spirit of the work was pleasant and earnest, and reduced to a minimum the element of boredom, probably the greatest influence in decreasing the quantity and quality of work in the ordinary routine of the school day.

The time schedule was so arranged as to test each class in the morning and in the afternoon. The classes were combined into groups of two, one being a half-grade below (A) or above (B) the other. The first class in a group took the first test in the morning of the first day and the second test in the afternoon of the second day following; the second class took the first test in the afternoon of the first day and the second test in the morning of the second day after. In this way the practice effect of the first class in the second test in the afternoon was balanced against that of the second class in the morning and consequently neutralized, leaving the fatigue effect free for comparison at the two periods. This arrangement was based upon a belief in the approximate equality of the practice effect in the two classes in a group. Even the slight differences that might have occurred were neutralized by giving the second test in the morning to the B classes and in the afternoon to the A classes of the first four groups, and then by giving the second test in the morning to the A classes and in the afternoon to the B classes of the second four groups.[1]

[1]One group was composed of 6A2 and 5B1, with the half-year relationship between classes as in the other groups but with different grades.

Two groups were tested in a day, the two classes of the first group at about 9:20 A. M. and 1:30 P. M., respectively, and the two classes of the second group at about 9:50 A. M. and 2:00 P. M., respectively. On the second day following, the classes tested before in the morning were tested in the afternoon, and *vice versa;* but the classes of the second group followed as before immediately after the corresponding classes of the first group, each group constituting a complete unit by itself with the same time relationship between the classes as in all the other groups. Testing two groups a day gave a wider representation of the school day, fifty minutes of morning work and fifty minutes of afternoon work.

The papers were graded by me by a gross method of counting answers only, very similar to that used by Mr. Courtis. Each test contained thirty-two points divided as follows: Two groups of two examples were each credited with one point, this point being counted wrong on account of one or more mistakes in either example; ten examples were each credited with one point, the point being counted wrong on account of one or more mistakes in the example; ten examples were each credited with two points, one point being counted wrong on account of one mistake, and two points being counted wrong on account of two or more mistakes, in the example. A gross method of counting answers over-emphasizes mistakes in proportion to the total amount of work done, not only at one period but even in a comparison between the percentage of error at two or more periods. The results of my New York tests showed that in the same papers the average increase in percentage of error at the three periods after the first period was by the gross method 3.2659 times that by the detailed method of counting every figure put down by a child in working the examples. The increase in percentage of error in the present results at the second period is, therefore, greater than the actual decrease in efficiency of the children tested. However, the gross method takes far less time for grading papers and is sufficiently accurate for comparison of the work at periods to which it has been uniformly applied.

Total Results.

School	Class	Pupils	FIRST PERIOD			SECOND PERIOD		
			Test	Right	Wrong	Test	Right	Wrong
Biggers ..	5A	43	1	12.67	5.02	2	16.53	6.14
" ..	5B	43	2	15.53	7.42	1	12.79	6.51
" ..	6A	44	1	15.61	7.34	2	16.64	9.23
" ..	6B	39	2	19.21	7.10	1	18.18	6.23
Monroe ..	5A	34	1	14.09	6.44	2	15.15	8.56
" ..	5B	46	2	15.04	7.26	1	10.96	6.39
" ..	6A	35	1	17.17	4.54	2	20.14	6.54
" ..	6B	32	2	21.00	6.62	1	16.59	5.94
Federal ..	7B	33	1	19.64	5.00	2	21.33	7.12
" ..	7A	41	2	19.27	7.12	1	16.98	6.66
" ..	6B	30	1	15.60	5.27	2	17.83	6.80
" ..	6A1	28	2	21.86	7.04	1	18.43	7.54
" ..	6A2	33	1	14.61	5.33	2	17.09	6.58
" ..	5B1	34	2	17.24	7.09	1	14.38	6.41
" ..	5B2	28	1	12.64	5.96	2	15.96	7.18
" ..	5A	30	2	14.40	6.37	1	11.47	6.60
	16	573		16.60	6.31		16.28	6.90
Sum of right and wrong points				22.91			23.18	
Relative amount done..				100.00			101.18	
Per cent right.....				72.46			70.23	
Relative per cent right..				100.00			96.92	

The average of right and wrong points for the sixteen classes at each period were added and the general average taken, in which all individual, class, and group variations were merged. *The sum of the general average of right and wrong points at the first period was compared with that at the second period, giving an increase of 1.18 per cent at the second period. The per cent right of the sum of right and wrong points at the first period was then compared with the per cent right at the second period, giving a decrease of 3.08 per cent at the second period.* The increase in quantity was probably due in part to the same influences as was the decrease in quality, greater speed and greater carelessness often going together.

Even the slight decrease in efficiency shown by these results is larger than the actual decrease in the classes tested, because

in Federal School the 5A class was somewhat disturbed in its afternoon work by the marching out of the other classes, and because the 7B and 6B classes had only a short indoor recess, on account of rain, before the afternoon tests. This is partly shown, though other influences may have entered, by the increase in quantity of 2.08 per cent and the decrease in quality of 3.37 per cent in the eight classes in Federal School, as compared with the increase in quantity of 0.22 per cent and the decrease in quality of 2.79 per cent in the eight classes in Biggers and Monroe Schools.

As the average per cent of increase in quantity in the New York tests, for the three periods after the first period, was by the detailed method of grading 1.0164 times that by the gross method, we can surmise that the 1.18 per cent increase in quantity in the Lynchburg tests by the gross method would be about 1.20 per cent by the detailed method. And as the average decrease in quality in the New York tests was by the detailed method only 0.3062 times that by the gross method, we can surmise that the 3.08 per cent decrease in quality in the Lynchburg tests would be about 0.94 per cent by the detailed method.

By a comparison of the results by the gross method in the twenty-five minute test in the afternoon in Lynchburg with the average of the two ten-minute tests in the afternoon in New York, we have 1.18 per cent increase in quantity in Lynchburg as compared with 1.79 per cent, and 3.08 per cent decrease in quality as compared with 5.61 per cent. In spite of the longer requirement for continuous work, the Lynchburg children show a smaller decrease in efficiency, probably on account of their better hygienic opportunities at home.

In order to determine whether the decrease in efficiency at the second period in Lynchburg was less in the more advanced grades, the following percentages were calculated: six fifth grades (three groups) showed an increase in quantity of 0.11 per cent and a decrease in quality of 2.82 per cent; six sixth grades (three groups) showed an increase in quantity of 0.36 per cent and a decrease in quality of 2.74 per cent; two seventh grades (one group) showed an increase in quantity of 2.08 per cent and a decrease in quality of 3.54 per cent. These slight differences seem to have no significance, the greater decrease in quality in the seventh grades being due mainly to the fact before mentioned that 7B had only a short indoor recess before the afternoon test.

An important consideration is the greater decrease in efficiency shown by the boys than by the girls, this difference even being noticeable in the greater restlessness of the boys during the tests. Complete tables of the sixteen class averages were made of the results by the boys and the girls, and the general averages, percentages, etc., were calculated. The boys showed an increase in quantity of 0.74 per cent and a decrease in quality of 4.25 per cent; the girls showed an increase in quantity of 1.62 per cent and a decrease in quality of 1.96 per cent. The boys showed an increase in quantity 0.4568 times that by the girls and a decrease in quality 2.1684 times that by the girls.

The practice effect was calculated by rearranging the class averages in columns according to tests rather than periods, thus neutralizing for the most part the decrease in efficiency at the second period. The sixteen classes showed an increase in quantity of 13.22 per cent in the second test and a decrease in quality of 0.50 per cent.

THIRD STUDY

The tests were given in December, 1914, to sixteen classes, 476 pupils, in the Intermediate School in Roanoke, Va. In this experiment reasoning problems in arithmetic were used in place of the four fundamental operations in the New York and Lynchburg experiments. Although the continued alertness required of grammar-grade pupils in rapid computation may well be considered indicative of hygienic efficiency for other school work, it must be admitted that the four fundamental operations may become partially automatic in well-drilled children, and, therefore, be less affected by fatigue. Furthermore, reasoning problems were used in the third experiment as involving processes and effort more generally needed in meeting school requirements.

In order to carry out in the third experiment the methods of the other two, it was necessary to use two reasoning tests that could be graded accurately and would be approximately equal in difficulty. The tests selected were Forms 1 and 3 of the Courtis Standard Test No. 8, since Mr. Courtis had proved by his very wide experience that these tests were as nearly equal as any so far made. They were given and scored according to the Courtis method, the number of examples done and the number right being calculated from the figures in the "Answer" column. The non-measureable elements in this gross method were mainly neutralized in the comparison of morning and afternoon work.

On Monday, December 8, Form 2 of Test No. 8 was given as a preliminary test to all sixteen classes in order to acquaint them with the matter and method of the succeeding tests; but the papers were not scored. By trying ten, fifteen, and twelve minutes as the time limit for this preliminary test, the twelve-minute limit was found to be most suitable for the classes concerned. For the succeeding tests the classes were divided into groups of four, each group containing a 7A, a 7B, a 6A, and a 6B class of equal rank with other classes of the same half-grade. As the classes in a group were tested in rapid succession, about sixty-five minutes were required for a group—from 9:25 to 10:30 in the morning or from 12:50 to 1:55 in the afternoon. The first test was given to Group 1 on Tuesday morning, to Group 2 on Tuesday afternoon, to Group 3 on Wednesday morning, and to Group 4 on Wednesday afternoon. The second test was given to Group 2 on Thursday morning, to Group 1 on Thursday afternoon, to Group 4 on Friday morning, and to

Group 3 on Friday afternoon. Thus each class was tested in the morning and in the afternoon; and each test was given in the morning and in the afternoon to the same number of classes of relatively equal rank. By this method the practice effect in the second test was neutralized by approximately equal representation in the morning and in the afternoon results. To complete this neutralization the classes in Groups 1 and 2 were tested in descending order (7A, 7B, 6A, 6B), and the classes in Groups 3 and 4 were tested in ascending order (6B, 6A, 7B, 7A).

A teacher generally remained in the room during a test, but assisted me only in distributing and collecting the papers. The children enjoyed the tests, which they thought were only for correctness and speed. They did not know the time limit for the work; and the few that finished early spent the extra time in looking over their papers. Where children were absent from one test or where they did not follow important directions, their papers were thrown out.

The hygienic conditions in the new school building were unusually good, but there had been no medical inspection of the children. Opening exercises were held in each room from 9:00 to 9:20, recess was given from 12:15 to 12:35, and the first bell for dismissal was rung at 2:15. The only recess did not give much invigoration, because the boys' playground was small and most girls remained indoors in spite of beautiful weather. Some pupils ate a light lunch at recess, but nearly all had dinner at home after school. The departmental system of instruction prevailed throughout with half-hour recitation periods. All classes included both boys and girls, the total represented in my results being 212 boys and 255 girls. The average age was 14.18 years.

After the papers were scored, the averages of the number of examples done and the number right in the morning and in the afternoon tests were made for each class. Then general averages were made for the sixteen classes combined, and the morning and the afternoon results were compared. *The number of examples done in the afternoon was 0.68 per cent greater than in the morning; the per cent of examples right in the afternoon was 3.22 per cent less than in the morning.* These per cents are strikingly similar to those from the Lynchburg and New York (average of two afternoon tests) experiments, where the increase in quantity in the afternoon was 1.18 and 1.79 per cent, respectively, and the decrease in quality 3.08 and 5.61 per cent,

respectively. As was shown by the New York results, the decrease in quality would be less by a detailed method of grading.

Total Results.

Class	Pupils	MORNING Test	MORNING Examples	MORNING Right	AFTERNOON Test	AFTERNOON Examples	AFTERNOON Right
7A1	26	1	5.96	4.50	2	6.73	5.27
7B1	30	1	5.60	4.20	2	5.97	4.77
6A1	31	1	4.87	3.81	2	5.45	4.23
6B1	29	1	4.83	3.76	2	5.07	3.97
7A2	32	2	6.84	5.44	1	5.78	4.00
7B2	35	2	6.23	5.09	1	5.97	4.31
6A2	27	2	6.37	4.81	1	5.93	4.19
6B2	25	2	5.88	4.56	1	4.96	3.00
6B3	24	1	4.96	2.71	2	5.33	3.58
6A3	26	1	5.04	3.27	2	6.42	4.19
7B3	35	1	4.74	3.34	2	6.06	4.51
7A3	33	1	6.64	4.70	2	7.48	5.27
6B4	28	2	6.00	4.43	1	5.39	3.54
6A4	29	2	5.79	3.83	1	5.10	3.31
7B4	28	2	6.82	4.93	1	5.96	4.00
7A4	29	2	7.00	5.55	1	6.72	5.14
16	467		5.85	4.31		5.89	4.20
Relative number done			100.00			100.68	
Per cent right			73.68			71.31	
Relative per cent right			100.00			96.78	

The practice effect was calculated by rearranging the class averages according to first or second test rather than according to morning and afternoon, thus neutralizing for the most part the decrease in efficiency in the afternoon. The sixteen classes showed in the second test an average increase of 12.43 per cent in quantity and an average increase of 7.16 per cent in quality. This large practice effect was mainly due to the difficulty the children had at first in disregarding unnecessary figures in some of the examples.

The seventh grade showed more relative efficiency in the afternoon than did the sixth grade, probably on account of better discipline in the former. The seventh grade had an increase of 1.61 per cent in quantity in the afternoon and a decrease of 2.82 per cent in quality; the sixth grade had a decrease of 0.04 per cent in quantity and a decrease of 3.67 per cent in quality.

Complete tables of class and general averages were made for the girls and for the boys, showing the greater relative efficiency of the former in the afternoon. The girls had an increase of 1.78 per cent in quantity in the afternoon and a decrease of 2.48 per cent in quality; the boys a decrease of 1.30 per cent in quantity and a decrease of 3.76 per cent in quality. These significant results are similar to those from the Lynchburg tests.

CONCLUSIONS

The most important features of these experiments were the giving (a) of actual school tests (b) under actual school conditions (c) by the same examiner with the same methods (d) to a large number of classes and pupils. Every possible effort was made to get the work and attitude of the routine school day and to eliminate individual and group variations. The comparison of morning and afternoon results probably neutralized the effect of two differences from routine conditions: the tests were given by an outsider and offered a break in the regular school requirements.

The per cent of increase in quantity and of decrease in quality in the afternoon results, as compared with the morning results, by the gross method of grading, may be summarized as follows:

City	Classes	Pupils	Increase in Quantity	Decrease in Quality
New York[1]	40	1,153	1.79	5.61
Lynchburg	16	573	1.18	3.08
Roanoke	16	476	0.68	3.22
	72	2,202	1.22	3.97

Seventy-two classes, 2,202 children, in three cities, showed in the afternoon results an average increase in quantity of 1.22 per cent and an average decrease in quality of 3.97 per cent.

The following tentative conclusions seem justified for grammar-grade pupils:[2]

1. Mental fatigue in relation to the daily school program is far less than is generally believed, even if the tendency to do a greater amount of work with a greater per cent of error can be partly attributed to fatigue.

[1]The averages of the two afternoon tests in New York are compared with the first morning test, as the second morning test did not represent the time or freshness of the morning test in Lynchburg and Roanoke, and the results were even below the afternoon average.

[2]These conclusions will probably not apply to primary grades (to most children under eleven years of age), without modification as to length of daily school program, etc.

2. The opinion regarding the graduated advantages of successive periods of the school day, and the sentiment in favor of putting arithmetic and other supposedly difficult subjects near the opening of the morning and afternoon sessions, with an emphasis upon the morning session, have not a sufficiently proved basis.

3. The decrease in quality of work by pupils as the day advances, considered to be more or less general in schools, is mainly due (a) to improper ventilation, lighting, heating, seating, etc., (b) to physical defects in the children, and (c) to loss of interest by pupil and teacher in the monotonous school routine.

4. With sound bodies, a hygienic school, proper classification, a vital and varied curriculum, and live teachers, most grammar-grade pupils will present no problem of fatigue in relation to the daily school program.

VITA.

William Harry Heck: Born Raleigh, N. C., 1879; B. A., Wake Forest College, N. C., 1897; Assistant Principal, Raleigh Male Academy, 1897-98; M. A., Wake Forest College, 1899; Graduate Student, Columbia University, 1899-1902; University Fellow, 1900-01; Honorary Fellow, 1901-02; Assistant Secretary, General Education Board, 1902-05; Professor of Education, University of Virginia, 1905—; Graduate Student, Columbia University, 1911-12.